JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 66D, No. 5, September–October 1962

Some Statistical Theory for the Analysis of Radio Propagation Data

M. M. Siddiqui

Contribution from Boulder Laboratories, National Bureau of Standards, Boulder, Colo.

(Received March 29, 1962)

The statistical theory of stationary processes has wide applications in the analysis of radio wave propagation data. In this paper, assuming the knowledge of the basic concepts of probability theory on the part of the reader, characteristics of stationary processes such as covariance and spectral density functions have been developed, problems of estimating these characteristics have been tackled, and numerous examples have been worked out to illustrate the theory.

1. Introduction

The indeterminacy in science, in microscopic as well as macrocopic physical processes, is gradually replacing the concept of "a cause and its effect" by the concept of "a cause and the probability distribution of its possible effects." We no longer ask the questions: will this signal be received? will a meteor be observed during this time interval? Instead we ask: what is the probability that this signal will be received? what is the probability that during this time interval at least one meteor will be observed?

In radio science the probability theory and statistical methods have wide applications, both in the development of theories and in the analysis of data. Scattering of radio waves from an irregular surface, propagation of electromagnetic field in an inhomogeneous medium, reflection of radio signals from meteor trails, the interaction of solar and cosmic energy with the earth's atmosphere, and many other such processes require theoretical models based on some kind of probability mechanism. Some models for these phenomena have alrady been proposed (to quote a few examples: Rayleigh [1899, 1919], Rice [1951], Isakovich [1952], Hoffman [1955, 1959], Wheelon [1960]), but essentially the field is wide open. The real challenge is to develop a sufficient number of stochastic models which would provide proper methods of analyzing mountains of data, which have already been collected, and point out ways for further meaningful experiments. In the absence of satisfactory theories, from which deductions can be made, we can only rely on *ad hoc* inferential hypotheses based on the statistical analysis of the data. Hence we turn to the question: what statistical methods can be used profitably to reduce the data into a few meaningful numbers? We find that the statistical theory of stationary processes, if correctly applied, has been and can be of great use. We shall, therefore, confine our study to the characteristics of stationary processes and efficient procedures for estimating these characteristics. The knowledge of the basic concepts of probability theory will be assumed on the part of the reader.

2. Definitions

A *random process*, $\{X(t)\}$, is a function of a parameter t, such that for each value of t, $X(t)$ is a random variable. We will refer to t as time, although it may be an arbitrary parameter. If t takes on only discrete values, say . . ., $-2, -1, 0, 1, 2,$ . . ., the process, $\{\ldots, X(-2), X(-1), X(0), X(1), X(2), \ldots\}$ is called a *discrete-time* process; and if t takes on values in a continuum, say $-\infty \leq t \leq \infty$, $\{X(t), -\infty \leq t \leq \infty\}$ is called a *continuous-time* process. Thus, for example, if we observe the hourly median value, $M(t)$, of the envelope, $X(t)$, of the received signal, $\{M(t)\}$ is a discrete-time and $\{X(t)\}$ is a continuous-time process. An observed record of $\{X(t)\}$, written without the curly brackets as $X(t)$, is called a *sample function*. The curly brackets simply denote that there are infinitely many possible sample functions which constitute the random process. If there is one and only one possible sample function, then the process is not random but deterministic.

What we want to know is the joint probability distribution of the random variables $X(t_1)$, . . ., $X(t_n)$ for arbitrary n, t_1, . . ., t_n. In its generality it is an impossible task, as it requires infinitely many sample functions observed over infinite time intervals. Thus we cannot proceed further with the analysis without assuming certain structure for the process.

A random process, $\{X(t)\}$, is called *weakly stationary* if, for all t and s,

$$EX^2(t) < \infty;\ EX(t) = EX(0);$$

$$EX(t)X(t+s) = EX(0)X(s).$$

Here, if X is a random variable with probability $(X \leq x) = P(x)$, EX stands for the mean value of X, i.e.,

$$EX = \int_{-\infty}^{\infty} x dP(x).$$

A random process, $\{X(t)\}$, is called *strictly stationary* if the joint distribution of $X(t_1+s), \ldots, X(t_n+s)$ is identical with the joint distribution of $X(t_1), \ldots, X(t_n)$ for every $n, t_1, \ldots, t_n$ and s.

A weakly or strictly stationary process, $\{X(t)\}$, is called *ergodic* in respect to a function g, if, for almost all sample functions, $X(t)$,

$$Eg(X(t)) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} g(X(t+s))ds.$$

If for any arbitrary function g, such that $Eg(X(t))$ exists, the relation given above holds, then the process is called, simply, *ergodic*.

3. Stationary Processes

In general, a record of some aspect of radio propagation cannot be considered as a sample function from a stationary process. For example, let $X(t)$, $t=1, 2, \ldots$, denote the hourly median value of the critical frequency, f_0F_2, of the F_2 layer of the ionosphere observed at Washington, D.C. From a priori considerations we would expect a diurnal, a seasonal, and a ten- or eleven-year cycle in the record, the last mentioned cycle corresponding to the sunspots cycle. However, if we eliminate these cycles by the least-squares fit we may reasonably assume that the residuals constitute a sample function of a stationary process. As we shall see later, theoretically it may be admissible to consider even a record with discernible periodicities in it as a sample function from a stationary process, but in a single record it is always advisable to remove such discernible cycles, and also the low-frequency part which, due to the limited extent of data, appears as a trend, before spectral analysis of the data. (For proper regression analysis refer to Siddiqui [1960].)

3.1. Continuous-Time Processes

Let $\{X(t)\}$ be a continuous-time real weakly stationary process with

$$EX(t)=0;\ EX(t)X(t+s)=\gamma(s)=\gamma(-s). \quad (3.1)$$

The question of estimating $EX(t)=\mu$, when μ is not assumed to be zero, will be taken up in the next section.

The spectral representation of a weakly stationary process with mean zero, suggested by Cramér [1940, 1942], is a powerful method of understanding the characteristics of the process.

If I_1 and I_2 are two disjoint intervals on a real line, a random set function $z(I)$ is called *orthogonal* if

$$Ez(I_1)\overline{z(I_2)}=0.$$

Here z denotes the complex conjugate of z. Let us write $dz(f)$ for $z(df)$. Cramér shows that $X(t)$ has the spectral representation

$$X(t) = \int_{-\infty}^{\infty} e^{2\pi i t f} dz(f). \quad (3.2)$$

where $z(f)$ is an orthogonal set function with

$$E|dz(f)|^2 = G(f+df) - G(f) = dG(f), df>0, \quad (3.3)$$

and $G(f)$, called *the spectral distribution function*, is a nondecreasing function with

$$G(-\infty)=0,\ G(\infty)=\gamma(0). \quad (3.4)$$

Now, $X(t)$ is real, hence

$$\begin{aligned}\gamma(s) &= EX(t)X(t+s) = E\overline{X(t)}X(t+s)\\ &= E\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{2\pi i(f-f')t+2\pi i f s}\, dz(f)\overline{dz(f')}\\ &= \int_{-\infty}^{\infty} e^{2\pi i f s} dG(f), \quad (3.5)\end{aligned}$$

from the orthogonality of $z(f)$. In (3.2) and (3.5) we have admitted negative as well as positive frequencies. It is possible to develop the theory in terms of positive frequencies only, in which case we will use cosine function for transformation rather than the exponential function. Obviously, exponential functions are much easier to work with than trigonometric functions. Hence, we will retain the representation as given above.

The integrals (3.2), (3.5), and others which will appear below should be interpreted in the Stieltjes sense. Thus, for instance, if $G(f)$ is a purely step function with steps of sizes $g_0', g_1', g_2', \ldots$, at $f=0, \pm f_1, \pm f_2, \ldots$, respectively, (3.5) is to be interpreted as

$$\begin{aligned}\gamma(s) &= g_0' + \sum_{k=1}^{\infty} g_k'(e^{2\pi i f_k s} + e^{-2\pi i f_k s})\\ &= g_0' + 2\sum_{k=1}^{\infty} g_k' \cos(2\pi f_k s). \quad (3.5a)\end{aligned}$$

g_k' is called *the spectral mass* at the frequencies $\pm f_k$. In this case

$$X(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos 2\pi f_k t + b_k \sin 2\pi f_k t), \quad (3.2a)$$

since $dz(f)$ is also a step function. Here

$$a_0 = dz(0),\ a_k = dz(f_k) + dz(-f_k),\ b_k = idz(f_k) - idz(-f_k).$$

Noting that $dz(-f_k) = \overline{dz(f_k)} = (a_k + ib_k)/2$, and remembering the orthogonality property of $z(f)$, and that $E\,X(t)=0$, we have

$Ea_k = Eb_k = 0$, all k; $Ea_k b_j = 0$, all k, j; $Ea_0^2 = g_0'$,

$$Ea_k^2 = Eb_k^2 = 2g_k', Ea_k a_j = Eb_k b_j = 0, k \neq j. \quad (3.6)$$

If, on the other hand, $G(f)$ is differentiable with $G'(f) = g(f)$, and $g(-f) = g(f)$, (3.5) becomes

$$\gamma(s) = \int_{-\infty}^{\infty} e^{2\pi i f s} g(f) df = 2\int_0^{\infty} \cos(2\pi f s) g(f) df. \quad (3.5b)$$

In this case, $G'(f) = g(f)$ is called *the spectral density function* of the process. More generally

$$G(f) = G_1(f) + G_2(f),$$

where $G_1(f)$ is purely a step function, and $G_2(f)$ admits the spectral density $g_2(f)$. Thus, the more general representations of $\gamma(s)$ and $X(t)$ are the following:

$$\gamma(s) = g_0' + \sum_{k=1}^{\infty} 2g_k' \cos(2\pi f_k s) + 2\int_0^{\infty} \cos(2\pi f s) g_2(f) df, \quad (3.5c)$$

$$X(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos 2\pi f_k t + b_k \sin 2\pi f_k t) + \int_{-\infty}^{\infty} e^{2\pi i f t} dz_2(f), \quad (3.2c)$$

where $E|dz_2(f)|^2 = g_2(f)df$: and a_0, a_k, b_k, not only satisfy (3.6) but also are uncorrelated with $dz_2(f)$. It may be noted that, since a_0, a_k, b_k are independent of t, in a single sample function they appear as constants; hence the desirability of estimating these constants by the least-squares method as discussed earlier in the section.

From here onwards we shall assume that by proper regression analysis $G_1(f)$ is eliminated so that $G(f)$ is differentiable, and the representation (3.5b) holds.

If

$$\int_{-\infty}^{\infty} |\gamma(s)| ds = 2\int_0^{\infty} |\gamma(s)| ds < \infty, \quad (3.7)$$

(3.5b) can be inverted to give

$$g(-f) = g(f) = \int_{-\infty}^{\infty} e^{-2\pi i f s} \gamma(s) ds = 2\int_0^{\infty} \cos(2\pi f s) \gamma(s) ds, \quad (3.8)$$

and g(f) *is continuous* everywhere. Thus *a sufficient condition for the continuity of* g(f) *is that* $|\gamma(s)|$ *be integrable.*

It is not sufficient, however, that only $\gamma(s)$ be integrable, for example, if

$$\gamma(s) = \frac{\sin 2\pi B s}{\pi s}, B > 0.$$

$$g(f) = 1, \text{ if } |f| \leq B; 0, \text{otherwise},$$

which is discontinuous at $f = \pm B$. It may be noted that, by definition, $|g(f)|$ is integrable. In fact $|g(f)| = g(f)$, and

$$\int_{-\infty}^{\infty} g(f) df = G(\infty) = \gamma(0) < \infty;$$

hence, *if a continuous-time process possesses a spectral density function, its covariance function* γ(s) *is continuous everywhere.*

Many a time we wish to obtain the relationship between the spectral density functions of two processes $\{X(t)\}$ and $\{Y(t)\}$, which are related through some linear differential or integral equation. Thus, for example, the frequency of a signal is the derivative of its phase.

Let $\{X(t)\}$ be a weakly stationary process, $W(t)$ a real integrable function, and let

$$Y(t) = \int_{-\infty}^{\infty} W(u) X(t-u) du = \int_{-\infty}^{\infty} W(t-u) X(u) du. \quad (3.9)$$

$W(t)$ is called *the linear filter*, $X(t)$ *the input*, and $Y(t)$ *the output* of the filter $W(t)$. If $W(t) = 0$ for $t < 0$, $W(t)$ is a physically realizable filter.

Let

$$X(t) = \int_{-\infty}^{\infty} e^{2\pi i t f} dz_x(f),$$

$$W^*(f) = \int_{-\infty}^{\infty} e^{-2\pi i u} W(u) du. \quad (3.10)$$

From (3.9), we obtain

$$Y(t) = \int_{-\infty}^{\infty} W^*(f) e^{2\pi i t f} dz_x(f).$$

Thus $Y(t)$ has the representation

$$Y(t) = \int_{-\infty}^{\infty} e^{2\pi i t f} dz_y(f),$$

where

$$dz_y(f) = W^*(f) dz_x(f). \quad (3.11)$$

Hence, by (3.3),

$$g_y(f) = |W^*(f)|^2 g_x(f). \quad (3.12)$$

where $g_y(f)$ and $g_x(f)$ are the spectral density functions of $\{Y(t)\}$ and $\{X(t)\}$, respectively. Thus $\{Y(t)\}$ is also a weakly stationary process. It is to be noted that the integrability of $W(t)$ is an essential condition for (3.12) to hold. Also, we require that $g_y(f)$ be integrable.

Let us now consider

$$U(t) = a_p X^{(p)}(t) + a_{p-1} X^{(p-1)}(t) + \ldots + a_0 X(t), \quad (3.13)$$

where $X^{(j)}(t)=d^jX/dt^j$. Again, using the spectral representation of $\{X(t)\}$ process, we obtain

$$dz_u(f)=(a_0+(2\pi if)a_1+\ldots+(2\pi if)^pa_p)dz_x(f)$$

$$g_u(f)=\left|\sum_{j=0}^{p}(2\pi if)^ja_j\right|^2 g_x(f). \qquad (3.14)$$

Of course, $\{U(t)\}$ will be defined if and only if $g_u(f)$ is integrable, i.e., when $\gamma_u(0)<\infty$.

A particular case of interest is obtained from (3.13) by setting $a_p=1$, $a_j=0$, $j\neq p$, so that $U(t)=X^{(p)}(t)$. In this case

$$g_u(f)=(2\pi)^{2p}f^{2p}g_x(f). \qquad (3.15)$$

Thus the process $\{X(t)\}$ *is differentiable* p *times if and only if the* 2p*th (hence every lower order) moment of the spectral density function,* $g_x(f)$, *exists.*

Differentiating $\gamma_x(t)$, $2p$ times in the relation

$$\gamma_x(t)=\int_{-\infty}^{\infty}e^{2\pi ift}g_x(f)df,$$

we obtain, from (3.15)

$$\gamma_u(t)=(-1)^p\gamma_x^{(2p)}(t). \qquad (3.16)$$

Thus, equivalently, the process $\{X(t)\}$ *will be differentiable* p *times if and only if* $\gamma_x(t)$ *is differentiable* 2p *times.*

Since for any process $\{X(t)\}$, $|\gamma_x(t)|\leq\gamma_x(0)$, $\gamma_x(t)$ attains its maximum at $t=0$. Hence, if $\gamma_x(t)$ is differentiable $2p$ times, we must have

$$0=\gamma_x^{(1)}(0)=\gamma_x^{(3)}(0)=\ldots=\gamma^{(2p-1)}(0), \qquad (3.17)$$

since each of

$$\gamma_x(t),\ -\gamma_x^{(2)}(t),\ \gamma_x^{(4)}(t),\ \ldots,\ (-1)^{p-1}\gamma_x^{(2p-2)}(t),$$

is a covariance function of a stationary process. Furthermore, $EX(t)X^{(j)}(t+s)=\gamma^{(j)}(s)$; hence $X(t)$ and $X^{(2j-1)}(t)$, $j=1,\ldots,p$, will be uncorrelated.

3.2. Discrete-Time Processes

It is a common practice to observe a process at equal intervals of time even though the process may be a continuous-time process. Let t be measured in seconds so that f is measured in cycles per second. Let the spacing between observations be h seconds, so that the derived process is $\{X(kh)\}$, $k=\ldots,-1,0,1,\ldots$, with covariance function $\gamma_k=\gamma(kh)$, $k=0,\pm1,\pm2,\ldots$. From (3.5b)

$$\gamma_k=\gamma(kh)=\int_{-\infty}^{\infty}e^{2\pi ifkh}g(f)df$$

$$=\sum_{r=-\infty}^{\infty}\int_{(2r-1)/(2h)}^{(2r+1)/(2h)}e^{2\pi ifkh}g(f)df$$

$$=h\int_{-1/(2h)}^{1/(2h)}e^{2\pi ifkh}g_1(f)df, \qquad (3.5d)$$

where

$$hg_1(f)=g(f)+\sum_{r=1}^{\infty}\left\{g\left(f+\frac{r}{h}\right)+g\left(f-\frac{r}{h}\right)\right\},\quad -1/(2h)\leq f\leq 1/(2h). \qquad (3.19)$$

The frequencies $f\pm r/h$, $r=1,2,\ldots$, which become indistinguishable from the frequency f, are called *aliases* to f. In case $\Sigma|\gamma_k|<\infty$, (3.5d) can be inverted to give

$$g_1(f)=\gamma_0+2\sum_{k=1}^{\infty}\gamma_k\cos(2\pi fkh), -1/(2h)\leq f\leq 1/(2h). \qquad (3.8d)$$

Of course, if we start with a discrete-time process, we may conveniently set $h=1$, and obtain

$$\gamma_k=2\int_0^{1/2}\cos(2\pi fk)g_1(f)df, k=0,\pm1,\pm2,\ldots,$$

$$g_1(f)=\gamma_0+2\sum_{k=1}^{\infty}\gamma_k\cos(2\pi fk),\ -1/2\leq f\leq 1/2.$$

Also, the spectral representation of $\{X(k)\}$ is

$$X(k)=\int_{-1/2}^{1/2}e^{2\pi ikf}dz_x(f),$$

with $E|dz_x(f)|^2=g_1(f)df$.

Relations corresponding to (3.12) and (3.14) can easily be obtained by replacing integrals with summations and differential equations with difference equations.

4. Estimation of the Mean

Let $\{X(t)\}$ be a continuous-time weakly stationary process with the mean μ, the covariance function $\gamma(s)$, and the spectral density function $g(f)$. Let a sample function $X(t)$, $0\leq t\leq T$, be available. Consider the sample mean

$$m=T^{-1}\int_0^T X(t)dt. \qquad (4.1)$$

We have $Em=\mu$, and [Siddiqui, 1961, eq (2.9)] the variance of m,

$$\text{var } m=2T^{-1}\int_0^T(1-s/T)\gamma(s)ds. \qquad (4.2)$$

The variance of m can also be expressed in terms of $g(f)$. In fact, using (3.5b) in (4.2) and interchanging the order of integration with respect to f and s, we obtain

$$\text{var } m=2\int_0^{\infty}\frac{\sin^2\pi Tf}{(\pi Tf)^2}g(f)df. \qquad (4.3)$$

Changing the variable of integration in (4.3) by

setting $f'=\pi Tf$, we have

$$\operatorname{var} m=\frac{2}{\pi T}\int_0^\infty \frac{\sin^2 f}{f^2}\, g\left(\frac{f}{\pi T}\right) df, \tag{4.4}$$

where again f' is replaced by f. Expressions (4.2) to (4.4) are exact. However, *if* g(f) *is continuous at* f=0 *and* g(0)>0, an asymptotic expression for var m is

$$\operatorname{var} m\simeq \frac{g(0)}{T}=\frac{2}{T}\int_0^\infty \gamma(s)ds. \tag{4.5}$$

Thus when $g(f)$ is continuous at zero and (*i*)$g(0)>0$, var $m=O(T^{-1})$, (*ii*)$g(0)=0$, var $m=0(T^{-1})$. In any case, var $m\to 0$ as $T\to\infty$ whenever $g(f)$ is continuous at zero, so that m tends to μ in probability (ergodic property).

For a discrete sample $X_k=X(kh)$, $k=1, 2, \ldots, N$, the corresponding results are

$$m=N^{-1}\sum_{i=1}^{N} X_k, \tag{4.1d}$$

$$\operatorname{var} m=\frac{\gamma_0}{N}+\frac{2}{N}\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\gamma_k \tag{4.2d}$$

$$\simeq\frac{g_1(0)}{N}. \tag{4.5d}$$

Here, $\gamma_k=\gamma(kh)$ and (4.5d) holds when $g_1(f)$ is continuous at zero.

Let N' be defined by the equation

$$N'=\gamma(0)/\operatorname{var} m, \tag{4.6}$$

where var m is given by (4.2) or (4.2d). N' will be called *the equivalent random sample size for estimating the mean*, as the variance of the mean of N' uncorrelated observations is $\gamma(0)/N'$.

Example 4.1. Let $\gamma(s)=\sigma^2e^{-\lambda|s|}\cos 2\pi f_0 s$, $\lambda>0$. This type of covariance function occurs sometimes in the analysis of radio propagation data. For a detailed discussion of it see chapter 5 of Bendat [1958]. Since $|\gamma(s)|$ is integrable, $g(f)$ is continuous everywhere. We have

$$g(0)=2\sigma^2\int_0^\infty e^{-\lambda s}\cos 2\pi f_0 s\, ds=\frac{2\sigma^2\lambda}{\lambda^2+4\pi^2f_0^2}.$$

Hence, for $\lambda T>>1$,

$$\operatorname{var} m\simeq\frac{2\lambda\sigma^2}{T(\lambda^2+4\pi^2f_0^2)}.$$

Of course the exact var m can be calculated from (4.2). If the sample is discrete, setting $\rho=e^{-\lambda h}$, where h is the sampling interval, we find

$$\operatorname{var} m\simeq\frac{\sigma^2(1-\rho^2)}{N(1+\rho^2-2\rho\cos 2\pi f_0\lambda h)}.$$

We may remark here that $\gamma(s)$ is not differentiable at $s=0$, hence the process is not differentiable.

Example 4.2. Let $\gamma(s)=\sigma^2e^{-\lambda|s|}$, $\lambda>0$. This is a special case of Example 4.1, when $f_0=0$. Thus

$$\operatorname{var} m\simeq\frac{2\sigma^2}{\lambda T}\ \text{or}\ \frac{\sigma^2(1+\rho)}{N(1-\rho)},$$

as the case may be. If $\lambda=0.69315$, and $h=1$, the $\rho=0.5$, and the equivalent random sample size N' is approximately $0.35T$ or $0.33N$, respectively. Again, $\gamma(s)$, hence the process, is not differentiable.

Example 4.3. Let $\gamma(s)=\sigma^2\cos 2\pi f_0 s$, $f_0\neq 0$. Note that this cannot be considered as a special case of Example 4.1, as $\gamma(s)$ is not integrable. This is a case when $g(f)$ does not exist. However, we can either use (4.2), or (4.3) with $g(f)df$ replaced by $dG(f)$, where $dG(f)=1/2\sigma^2$, if $f=\pm f_0$; 0, otherwise. Thus

$$\operatorname{var} m=\sigma^2\frac{\sin^2\pi Tf_0}{(\pi Tf_0)^2}.$$

Thus var $m=O(T^{-2})$ instead of $O(T^{-1})$. Furthermore, if $T=k/f_0$, k a positive integer, var $m=0$.

Example 4.4. Let

$$g(f)=\frac{\sigma^2(1-a)}{2B^{1-a}}\frac{1}{|f|^a},\ 0<a<1, \text{ if } |f|\leq B; =0, \text{ if } |f|>B.$$

This type of spectral density function has been observed for the frequency fluctuations of the received signal, when the transmitted signal has constant frequency. Since $g(f)$ is specified instead of $\gamma(s)$ it is more convenient to use (4.3). Also, since $g(f)$ is not continuous at $f=0$, asymptotic approximation cannot be used. Setting $\pi TB=A$, we find

$$\operatorname{var} m=\frac{\sigma^2(1-a)}{A^{1-a}}\int_0^A\frac{\sin^2 x}{x^{2+a}}dx.$$

The integral can be evaluated numerically. An upper bound to var m is obtained by dominating $\sin^2 x$ by x^2 when $0\leq x\leq 1$, and by 1 when $x>1$. Thus, if $A\leq 1$, var $m<\sigma^2$; and, if $A>1$,

$$\operatorname{var} m<\frac{2\sigma^2}{(1+a)A^{1-a}}-\frac{\sigma^2(1-a)}{(1+a)A^2}.$$

Thus var $m=O(T^{-1+a})$ rather than $O(T^{-1})$. Note that $f^{2p}\,g(f)$ is integrable for $p=1, 2, \ldots$, hence the process is differentiable to any order.

5. Estimation of the Covariance and the Spectral Density Functions

In this section we will confine ourselves to discrete-time Gaussian processes.

Let $\{X_k\}$, $k=\ldots, -1, 0, 1, \ldots$, be a discrete-time Gaussian stationary process with mean μ, covariance function γ_k, and spectral density $g(f)$. $g(f)$ and γ_k are related by the transform pair

$$\gamma_k=2\int_0^{\frac{1}{2}}\cos(2\pi fk)g(f)df,\ k=0,\pm 1,\pm 2,\ldots,$$

640618—62——6

$$g(f)=\gamma_0+2\sum_{k=1}^{\infty}\gamma_k \cos(2\pi f k), -1/2\leq f\leq 1/2. \quad (5.1)$$

Let N consecutive observations be made on the process. We may denote them as $X_1, X_2, \ldots, X_N$. From the sample we calculate

$$m=N^{-1}\sum_{k=1}^{N}X_k, C_s=(N-s)^{-1}\sum_{k=1}^{N-s}(X_k-m)(X_{k+s}-m), \quad (5.2)$$

for $s=0, 1, 2, \ldots, n$, where $n\to\infty$ as $N\to\infty$ but $n/N\to 0$. Thus, for example, n may be the largest integer in $N^{1/2}$. If

$$\sum_{k=1}^{\infty}\gamma_k=a\gamma_0<\infty; \sum_{k=1}^{\infty}\gamma_k^2=b\gamma_0^2<\infty;$$

it can be shown that

$$EC_s\cong\gamma_s-\frac{\gamma_0(1+2a)}{N}, \quad (5.3)$$

$$\operatorname{var} C_s\cong N^{-1}\left[\gamma_0^2+\gamma_s^2+2\sum_{k=1}^{\infty}(\gamma_k^2+\gamma_{k-s}\gamma_{k+s})\right]. \quad (5.4)$$

Thus C_s is biased as an estimate of γ_s but the bias, and also the variance, tend to zero as $N\to\infty$. Thus $C_s\to\gamma_s$ in probability (ergodic property). We note that if $|\gamma_k|$ is summable, both γ_k and γ_k^2 will be summable (since $\gamma_k^2<\gamma_0|\gamma_k|$), and the ergodic property will hold.

We note in particular that the estimate of the variance γ_0 is C_0. From (5.4)

$$\operatorname{var} C_0\cong\frac{2\gamma_0^2(1+2b)}{N}, b=\sum_{1}^{\infty}\gamma_k^2/\gamma_0^2. \quad (5.5)$$

The *equivalent number of degrees of freedom for estimating the variance* may be defined as

$$N'=\frac{2(EC_0)^2}{\operatorname{var} C_0}\cong N\frac{\gamma_0^2\left[1-\frac{1+2a}{N}\right]^2}{\gamma_0^2(1+2b)}\cong\frac{N}{(1+2b)}.$$

Example 5.1. Let $\gamma_k=\sigma^2\rho^{|k|}$, $|\rho|<1$. Then $b=\rho^2/(1-\rho^2)$,

$$\operatorname{var} C_0\cong\frac{2\sigma^4(1+\rho^2)}{N(1-\rho^2)},$$

and

$$N'\cong\frac{N(1-\rho^2)}{1+\rho^2}.$$

Thus, if $\rho=\pm 0.5$, $N'\cong 3N/5$; if $\rho=\pm 0.9$, $N'\cong N/10$.

Example 5.2. Let $\gamma_k=\sigma^2 e^{-\lambda k^2}$, $\lambda>0$. Then

$$1+2b=\sum_{-\infty}^{\infty}e^{-2\lambda k^2}\cong[\pi/(2\lambda)]^{1/2}, \text{ if } \lambda<1,$$

$$\operatorname{var} C_0\cong\frac{2\sigma^4}{N}\left(\frac{\pi}{2\lambda}\right)^{1/2},$$

$$N'=N\left(\frac{2\lambda}{\pi}\right)^{1/2}.$$

Thus, if $e^{-\lambda}=0.5$, $N'\cong 0.66N$; if $e^{-\lambda}=0.9$, $N'\cong 0.26N$.

We turn now to the estimation of the spectral density function. The classical harmonic analysis leads to the periodogram estimate,

$$g_N\left(\frac{j}{N}\right)=\frac{N}{4}(a_j^2+b_j^2), j=1, 2, \ldots, N-1,$$

where N is assumed to be an odd integer, and where

$$a_j=\frac{2}{N}\sum_{k=1}^{N}X_k\cos\frac{2\pi jk}{N}; b_j=\frac{2}{N}\sum_{k=1}^{N}X_k\sin\frac{2\pi jk}{N}.$$

It can be shown that if $j/N\to f$ as $N\to\infty$, $\lim_{N\to\infty} Eg_N(j/N)=g(f)$, $\operatorname{var} g_N(f)\cong g^2(f)$; so that $g_N(f)$ does not converge to $g(f)$ as $N\to\infty$. We mention here two alternative estimates which converge in probability to $g(f)$, i.e., which are *consistent* estimates of $g(f)$.

(1) Bartlett [1950].

$$g_N^{(1)}(f)=\sum_{k=-n}^{n}\left(1-\frac{|k|}{n}\right)C_k\cos 2\pi kf,$$

$$f=\frac{j}{n}, j=0, 1, \ldots, n-1,$$

where $n\to\infty$ as $N\to\infty$ but $n/N\to 0$. It is found that

$$\lim_{N\to\infty} Eg_N^{(1)}(f)=g(f),$$

$$\operatorname{var} g_N^{(1)}(f)\cong\begin{cases}\frac{2n}{3N}g^2(f), f\neq 0,\\ \frac{4n}{3N}g^2(0), f=0.\end{cases}$$

(2) Blackman and Tukey [1958].

$$g_N^{(2)}(f)=C_0+2\sum_{k=1}^{n-1}\left(0.46\cos\frac{\pi k}{n}+0.54\right)C_k\cos\pi kf$$
$$+(0.46\cos\pi+0.54)C_n$$
$$\cos\pi nf, f=j/n, j=0, 1, \ldots, n-1;$$

$$\lim_{N\to\infty} Eg_N^{(2)}(f)=g(f),$$

$$\operatorname{var} g_N^{(2)}(f)\cong\begin{cases}0.8\frac{n}{N}g^2(f), \text{ when } f\neq 0,\\ 1.6\frac{n}{N}g^2(0), \text{ when } f=0.\end{cases}$$

Parzen [1957] discusses a general method of obtaining consistent estimates of the spectral density function.

Another powerful technique is to find a linear filter $W(t)$ such that if X_t is the input, the output, Y_t, is approximately a white noise, i.e., $g_y(f)=\sigma_y^2$, $-\frac{1}{2}\leq f\leq\frac{1}{2}$. The variance, σ_y^2, of Y_t can be estimated consistently, hence, $g_x(f)=|W^*(f)|^2\sigma_y^2$, can also be estimated consistently.

Example 5.3. Consider

$$X_t=a_0Y_t+a_1Y_{t-1}+\ldots+a_pY_{t-p}.$$

Here,

$$W(t)=a_t, \text{ if } t=0,1,\ldots,p=0, \text{ otherwise.}$$

Replacing integrals by summations in (3.9) and (3.10), we obtain

$$g_x(f)=\left|\sum_{t=0}^{p} a_te^{-2\pi itf}\right|^2 g_y(f).$$

If Y_t is a white noise, X_t is called *a moving average process*. On the other hand, if X_t is a white noise and

$$\sum_{k=0}^{p} a_kZ^{p-k}=0$$

has all the roots within the unit circle $|Z|=1$ in the complex plane, Y_t is a stationary process and is called an *autoregressive* process of order p. In the former case $g_y(f)=\sigma_y^2$, and in the latter case $g_x(f)=\sigma_x^2$.

Special case (a). If $a_k=1/(p+1)$, $k=0, 1, \ldots, p$, and Y_t a white noise, X_t is called a *simple moving average process*. In this case $g_x(f)$ is evaluated to be

$$g_x(f)=\frac{1}{p+1}\frac{\sin^2\{(p+1)\pi f\}}{\sin^2\{\pi f)}\sigma_x^2, -\tfrac{1}{2}\leq f\leq\tfrac{1}{2};$$

as $\sigma_y^2=(p+1)\sigma_x^2$. We only need a consistent estimate of σ_x^2, which is the sample variance.

Special case (b). $Y_t-\rho Y_{t-1}=X_t$, $|\rho|\leq 1$, and X_t a white noise. This is a first-order autoregressive scheme. Setting $a_0=1$, $a_1=-\rho$, and $a_t=0$ otherwise, we find

$$g_y(f)=(1-2\rho\cos 2\pi f+\rho^2)^{-1}g_x(f).$$

Also

$$g_x(f)=\sigma_x^2=\sigma_y^2(1-\rho^2);$$

hence

$$g_y(f)=\sigma_y^2(1-\rho^2)(1-2\rho\cos 2\pi f+\rho^2)^{-1}.$$

σ_y^2 and ρ are consistently estimated by the sample variance and the first (lag 1) autocorrelation ($=C_1/C_0$), respectively. Note that the covariance of $\{Y_t\}$, $\gamma_y(k)=\sigma_y^2\rho^{|k|}$. Hence, $g_y(f)$ can also be obtained directly from

$$g_y(f)=\sigma_y^2\left[1+\sum_{k=1}^{\infty}(\rho^ke^{2\pi ifk}+\rho^ke^{-\pi ifk})\right].$$

Example 5.4. In figure 1, 236 noon hour monthly median f_0F_2 values in Mc/s observed at Washington, D.C., from May 1934 to December 1953 are plotted against t. The unit of t is one month and $t=1$ corresponds to January 1934. Hence, the first value plotted corresponds to $t=5$. In figure 2 the autocorrelation function, $r(\tau)=C\tau/C_0$, of this data is plotted for $\tau=1, 2, \ldots, 120$; and in figure 3, the Blackman-Tukey spectral density function, $g_N^{(2)}(f)$, is graphed against frequency, f, cycles per year. $g_N^{(2)}(f)$ is the normalized density per cycle per year. The spectral density indicates that there are two fundamental cycles in the data corresponding to $f=0.1$ (10-yr cycle), and $f=1$ (one yr cycle). Besides these, their first two harmonics ($f=0.2$, 0.3, and $f=2$, 3), and their "interaction" frequencies ($f=1\pm 0.1=0.9$, 1.1) are also significant. Since the least common period for all these cycles is 120 months, before further analysis it seemed advisable to add four more terms to the data in figure 1, to make the total number of data points $240=2\times 120$. These values are 5.9, 5.7, 5.5, and 5.0 corresponding to the months of January–April 1954.

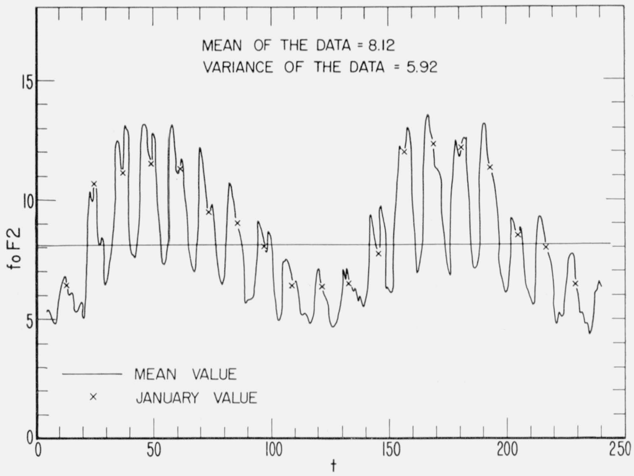

FIGURE 1. *Noon-hour monthly median f_0F_2 (Mc/s) at Washington, D.C., from May 1934 to December 1953.*

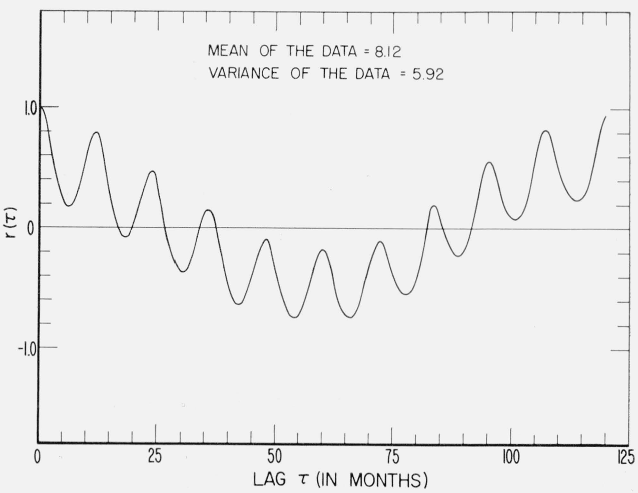

FIGURE 2. *Autocorrelation function of the data in figure 1.*

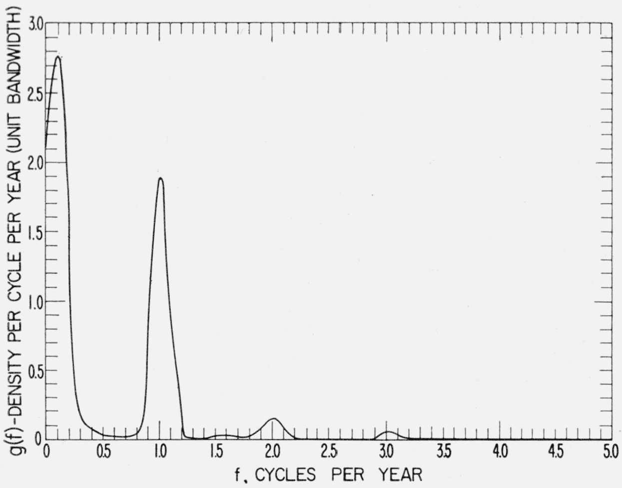

FIGURE 3. *Spectral density function of the data in figure 1.*

Let $x(t)$ denote the f_0F_2 value at time t. Then $x(t)$ is represented as

$$x(t)=m(t)+z(t),$$

where

$$m(t)=\bar{X}+\sum_j \left(a_j \cos \frac{2\pi jt}{120}+b_j \sin \frac{2\pi jt}{120}\right),$$

$$\bar{X}=N^{-1}\sum_t x(t),$$

and the summation over j is on $j=1, 2, 3, 9, 10, 11, 20$, and 30. $z(t)$ represents the "error" or "noise." The coefficients a_j and b_j are obtained by the least-squares method, and their values are as follows:

TABLE 1. *a_j, b_j, and $A_j^2+b_j^2$*

j	a_j	b_j	$a_j^2+b_j^2$
1	−2.24	1.17	6.38
2	−0.371	−0.297	0.226
3	−.0172	−.220	.0485
9	−.130	.393	.171
10	−.671	−1.77	3.58
11	.342	0.443	0.313
20	.556	.0556	.313
30	−.233	−.247	.115

After fitting $m(t)$, the residuals $z(t)$ are calculated from $z(t)=x(t)-m(t)$. Autocorrelation analysis of $z(t)$ shows that they can be considered as "white noise." To test whether the residuals are normally distributed, the range $(-\infty, \infty)$ is divided into 10 intervals such that in each interval the expected frequency is the same, i.e., 24. For this purpose an estimate of var z is required. This estimate is

$$s_z^2=\frac{\Sigma z^2(t)}{N-17}=\frac{87.1}{223}=0.39.$$

We note that $\Sigma z(t)=0$. The resulting class intervals and the observed frequencies, f_0, are tabulated in table 2.

TABLE 2

Class interval			f_0
$-\infty$	. . .	−0.801	23
−0.801	. . .	−.526	20
−.526	. . .	−.327	25
−.327	. . .	−.158	24
−.158	. . .	0	26
0	. . .	.158	24
0.158	. . .	0.327	33
.327	. . .	.526	20
.526	. . .	.801	20
.801	. . .	$.\infty$	25

The expected frequency, f_e, for each class interval is 24. Hence $\chi^2=\Sigma \frac{(f_0-f_e)^2}{f_e}=\frac{1}{24}\Sigma f_0^2-N$. This value is calculated to be 5.67, and the number of degrees of freedom for χ^2 is 8. The probability that such a sample or worse comes from a normal distribution, as judged by the χ^2 value, is more than 60 percent. We may thus conclude that we have essentially completed our analysis and that $m(t)$ is the best fit to the data.

The correlation coefficient between $x(t)$ and $m(t)$ is given by

$$R=\left(1-\frac{s_z^2}{s_x^2}\right)^{1/2}=\left(1-\frac{0.39}{5.96}\right)^{1/2}\cong 0.97.$$

$m(t)$ is plotted against t in figure 4 for $t=1, 2, \ldots, 120$. The values of $m(t)$ for $t=1, 2, \ldots, 12$ may be taken as predictions for the successive months of a year ending in 4, i.e., 1954, 1964, 1974, the values for $t=13$ to 24 for the months of a year ending in 5; and so on. Since the residual standard deviation is 0.62, uniform 95 percent confidence limits for $x(t)$ are $m(t)\pm 1.2$.

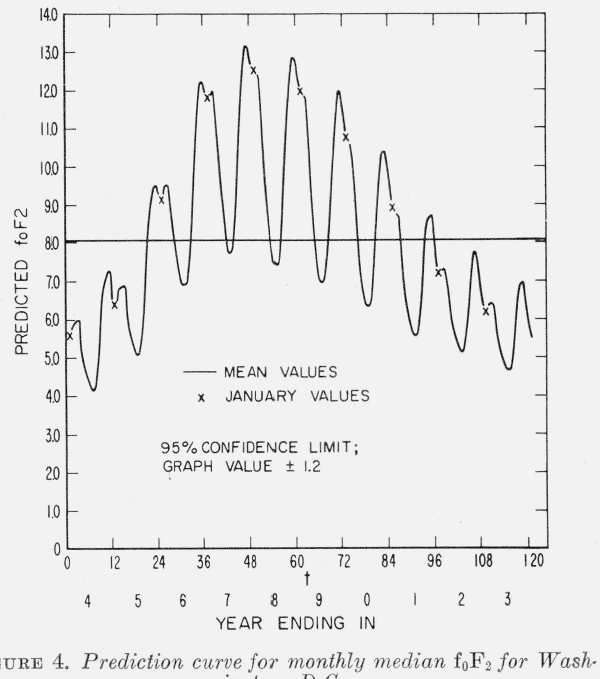

FIGURE 4. *Prediction curve for monthly median* f_0F_2 *for Washington, D.C.*

6. Estimation of the Distribution Function

The process distribution function, $P(x)$, is the probability that $X_t \leq x$. The sample distribution function, $P^*(x)$, is the proportion of $X_1, \ldots, X_N$, such that $X_t \leq x$. Thus, with $i=1, 2, \ldots, N$,

$$P^*(x)=\begin{cases} 0, \text{ if } x<\text{all } X_i, \\ j/N, \text{ if exactly } j \text{ of } X_i \leq x, \\ 1, \text{ if all } X_i \leq x. \end{cases} \tag{6.1}$$

We will assume $\{X_t\}$ to be strictly stationary process and introduce

$$\begin{aligned} P_s(x_1,x_2)&=Pr(X_t \leq x_1, X_{t+s} \leq x_2) \\ &=Pr(X_0 \leq x_1, X_s \leq x_2). \end{aligned} \tag{6.2}$$

Note that, if $x_2 \geq x_1$

$$\begin{aligned} P_0(x_1,x_2)&=Pr(X_t \leq x_1, X_t \leq x_2) \\ &=Pr(X_t \leq x_1)=P(x_1). \end{aligned} \tag{6.3}$$

Let

$$Y_t=1, \text{ if } X_t \leq x; \ 0, \text{ if } X_t > x; \tag{6.4}$$

then

$$P^*(x)=N^{-1}\sum_{t=1}^{N} Y_t; \tag{6.5}$$

i.e., $P^*(x)$ is a sample mean of the $\{Y_t\}$ process and the results of section 3 will apply, noting that

$$\begin{aligned} EY_t&=Pr(X_t \leq x)=P(x) \\ \gamma_y(s)&=E[Y_t-P(x)][Y_{t+s}-P(x)] \\ &=Pr(X_t \leq x, X_{t+s} \leq x)-P^2(x) \\ &=P_s(x,x)-P^2(x). \end{aligned} \tag{6.6}$$

Since $P_0(x,x)=P(x)$, we have

$$\gamma_y(0)=P(x)[1-P(x)]. \tag{6.7}$$

Thus, from (4.2d),

$$\operatorname{var} P^*(x)=\frac{\gamma_y(0)}{N}+\frac{2}{N}\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\gamma_y(k),$$

where $\gamma_y(k)$ are given in (6.6) and (6.7). It is obvious that if

$$\sum_{k=1}^{\infty}\gamma_y(k)=\sum_{k=1}^{\infty}[P_k(x,x)-P^2(x)]<\infty,$$

then $\operatorname{var} P^*(x) \to 0$ as $N \to \infty$, and $P^*(x)$ converges in probability to $P(x)$ (ergodic property).

To estimate $\operatorname{var} P^*(x)$ from the sample, we need an estimate of $P_k(x,x)$. For this purpose we find the proportion of the sample pairs (X_t, X_{t+k}) such that *both* $X_t \leq x$ and $X_{t+k} \leq x$. We will denote this proportion as $P_k^*(x,x)$.

Example 6.1. The following is a systematic sample of 80 observations (read at 5-sec intervals) of received field intensity in (microvolts)^2. Read left to right.

0.20,	0.71,	0.06,	0.05,	0.76,	0.32,	0.96,
0.63,	0.09,	0.18.	0.25,	0.45,	0.26,	0.10,
0.95,	0.01,	0.50,	1.20,	1.99,	0.32,	0.51,
0.01,	0.16,	0.56,	3.16,	1.27,	2.24,	1.00,
0.81,	1.29,	0.28,	0.21,	0.35,	0.20,	0.39,
0.89,	1.24,	0.08,	0.98,	1.01,	0.49,	0.90,
1.90,	1.42,	1.56,	1.32,	1.20,	1.59,	2.40,
2.24,	0.80,	0.56,	1.45,	0.18,	0.02,	0.28,
0.81,	0.18,	1.13,	0.64,	1.95,	0.48,	0.55,
0.44,	0.28,	0.07,	0.71,	0.48,	0.40,	0.06,
0.79,	1.01,	0.51,	0.70,	0.14,	0.16,	0.01,
0.06,	0.03,	0.01.				

Let us consider the estimate of $P(0.5)$. We obtain

$$P^*(0.5)=\frac{39}{80}=0.4875;$$

$$P_1^*(0.5, 0.5)=\frac{23}{79}=0.2911, P_2^*(0.5, 0.5)=\frac{21}{78}=0.2692;$$

$$P_3^*(0.5, 0.5)=\frac{20}{77}=0.2597, P_4^*(0.5, 0.5)=\frac{16}{76}=0.2105.$$

Noting that

$$P^{2*}(0.5)=0.2377, \ \hat{\gamma}_y(0)=P^*(0.5)[1-P^*(0.5)]=0.2498,$$

where $\hat{\gamma}_y(k)$ denotes the sample estimate of $\gamma_y(k)$, we find that $\hat{\gamma}_y(3)$ and $\hat{\gamma}_y(4)$ are of opposite sign and $\hat{\gamma}_y(3)/\hat{\gamma}_y(0)$, $\hat{\gamma}_y(4)/\hat{\gamma}_y(0)$ are negligible compared to unity. We may assume that $\sum_3^N \gamma_y(k)$ is negligible. Thus

$$\begin{aligned} \operatorname{var} P^*(0.5) &\cong \frac{0.2498}{80}+\frac{2}{80}\left[\left(1-\frac{1}{80}\right)(0.0534) + \left(1-\frac{2}{80}\right)0.0315\right] \\ &=0.0031+0.0021 \\ &=0.0052. \end{aligned}$$

With the assumption of approximate normality for the distribution of $P(0.5)$, 95 percent confidence limits for $P(0.5)$ are

$$P^*(0.5)-1.96s_p \leq P(0.5) \leq P^*(0.5)+1.96s_p,$$

where

$$s_P=\sqrt{\operatorname{var} P^*(0.5)}=0.072, \text{ i.e., } 0.35 \leq P(0.5) \leq 0.63.$$

If we assume the Rayleigh power distribution for the above data, the distribution function is most efficiently estimated as

$$\hat{P}(x)=1-e^{-\frac{x}{0.71}}, 0\leq x\leq \infty,$$

where 0.71 is the mean of the observations. Thus, under the assumption of the Rayleigh distribution, $P(0.5)$ is estimated to be

$$\hat{P}(0.5)=1-e^{-\frac{0.50}{0.71}}=0.51.$$

7. References

Bartlett, M. S., Periodogram analysis and continuous spectra, Biometrika **37,** 1–16 (1950).

Bendat, J. S., Principles and Applications of Random Noise Theory (John Wiley & Sons, Inc., 1958).

Blackman, R. B., and J. W. Tukey, The measurement of power spectra from the point of view of communications engineering, Parts I and II, The Bell System Technical Journal **37,** Nos. 1, 2, pp. 185–282, 485–569 (1958).

Cramér, H., On the theory of stationary random processes, Ann. Math. **41,** 215–230 (1940).

Cramér, H., On harmonic analysis in certain functional spaces, Ark. Mat. Astr. Fys. **28B,** 1–7 (1942).

Hoffman, W. C., Scattering of electronmagnetic waves from a random surface, Quarterly of Applied Mathematics **13,** No. 3, 291–304 (1955).

Hoffman, W. C., The electromagnetic field in a randomly inhomogeneous medium, IRE Trans. Antennas and Propagation **AP–7,** Special Supplement, S301–S306 (1959).

Isakovich, M. A., The scattering of waves from a statistically rough surface, Zhurnal Eksp. Teorl. Fiziki **23,** 305–314 (1952).

Parzen, E., On consistent estimates of the spectrum of a stationary time series, Ann. Math. Statistics **28,** 329–348 (1957).

Rayleigh, Lord, On the problem of random vibrations and of random flights in 1, 2, and 3 dimensions, Scientific Papers **1** (Cambridge University Press, 1899); Phil. Mag. **37,** 321–347 (1919).

Rice, S. O., Reflections of electromagnetic waves from slightly rough surfaces, Communs. on Pure and Appl. Math. **4,** 351–378 (1951).

Siddiqui, M. M., Tests for regression coefficients when errors are correlated, Annals of Mathematical Statistics **31,** 929–938 (1960).

Siddiqui, M. M., Some properties of the empirical distribution function of a random process, J. Research NBS **65B** (Math. and Math. Phys.), 117–127 (1961).

Wheelon, A. D., Amplitude distribution for radio signals reflected by meteor trails, J. Research NBS **64D** (Radio Prop.), 449–454 (1960).

(Paper 66D5–217)